

Exploring Sounded and Silent Laughter in Multiparty Task-based and Social Interaction - Audio, Video and Biometric Signals

Emer Gilmartin¹, Shannon Hennig², Ryad Chellali², Nick Campbell¹

¹ Speech Communications Lab, Trinity College Dublin

² PAVIS Group, Istituto Italiano di Tecnologia

gilmare@tcd.ie

Abstract

We report on our explorations of laughter in multiparty spoken interaction. Laughter is universally observed in human interaction. It is multimodal in nature: a stereotyped exhalation from the mouth in conjunction with rhythmic head and body movement. Predominantly occurring in company rather than solo, it is believed to aid social bonding. Spoken interaction is widely studied through corpus analysis, often concentrating on ‘task-based’ interactions such as information gap activities and real or staged business meetings. Task-based interactions rely heavily on verbal information exchange while the immediate task in natural conversation is social bonding. We investigate laughter *in situ* in task-based and social interaction, using corpora of non-scripted (spontaneous) multiparty interaction: the task-oriented AMI meetings corpus, and the conversational TableTalk, and D-ANS corpora. We outline extension of previous work on laughter and topic change, describe the collection of natural social spoken interaction in the D-ANS corpus including audio-visual and biosignals, and describe an annotation experiment on multimodal aspects of laughter. We discuss the results and signal current and future research directions.

Keywords: laughter, conversation analysis, social conversation, multimodal communication

1 Introduction

In our work, we investigate natural face-to-face social and task-based spoken interaction. Human conversational interaction is a multi-faceted, multi-modal, and multi-functional activity, in which participants filter information from a bundle of signals and cues, many of them temporally interdependent. From this information, the interlocutors make inferences on the speaker’s literal and pragmatic meaning, intentions, and affective state. This paper reports some of our recent work on laughter, an intrinsic thread of the conversational information bundle. Below we outline the background to our work, describe the corpora we use in our investigations, report problems encountered with laughter annotations in earlier work and describe an experiment in the multimodal annotation of laughter and the replication of an earlier study on laughter around topic change.

1.1 Task-based and Social Conversation

Communication is situated, and its characteristics vary with the type of interaction or ‘speech-exchange system’ (Sacks, Schegloff, and Jefferson 1974) participants are engaged in. While types or genres of written communication are fairly clear, categorization is not regarded as straightforward for speech genres. The range of interactions involving speech which humans engage in is enormous, with the problem of categorizing different types of speech exchange

into genres labeled as ‘notorious’ (Bakhtine 1986). Although there are many genres of spoken interaction, much attention has been paid to ‘task-based’ or ‘transactional’ dialogue. Indeed, dialogue system technology is based on task-based dialogue (Allen et al. 2001) for reasons of tractability. However, in real-life conversation there is often no obvious short term task to be accomplished through speech and the purpose of the interaction is better described as a short or longer term goal of building and maintaining social bonds (Malinowski 1923; Dunbar 1998). In these situations, the transfer of verbal information is not the main goal, and non-verbal information may carry more importance in attaining the goals of social bonding. Conversation is ‘interactional’ rather than ‘transactional’. As an example, a tenant’s short chat about the weather with the concierge of an apartment block is not intended to transfer important meteorological data but rather to build a relationship which may serve either of the participants in the future. The study of different types of interaction, and indeed stretches of social and task-based communication within the same interaction sessions, can help discriminate which phenomena in one type of speech exchange system are generalizable and which are situation or genre dependent. In addition to extending understanding of human communication, such knowledge will be useful in human machine interaction design, particularly in the field of ‘companion’ robots or ‘relational agents’, as the very notion of a companion application entails understanding of social spoken interaction (Gilmartin and Campbell 2014). As part of our larger exploration of how aspects of interaction vary in social (chat) and task-based scenarios, we are investigating the occurrence and role of laughter in multiparty interaction.

1.2 Laughter in Spoken Interaction

Laughter is a universally observed element of human interaction, part of the gesture call system (Burling 2007), appearing in human behavior before language both in evolutionary terms and in child development, and believed to have

evolved from the primate social bonding (Glenn 2003). Laughter is predominantly a shared rather than solo activity, with most incidences reported as occurring in the company of others (Provine 2004). Laughter occurs widely in spoken interaction, but punctuates rather than interrupts speech (Provine 1993). Laughter episodes take a range of forms – from loud bouts to short, often quiet chuckles. Laughter is described as a stereotyped exhalation of air from the mouth in conjunction with rhythmic head and body movement (Mehu and Dunbar 2008), and is thus multimodal, with visual and acoustic elements available to the interlocutor’s senses.

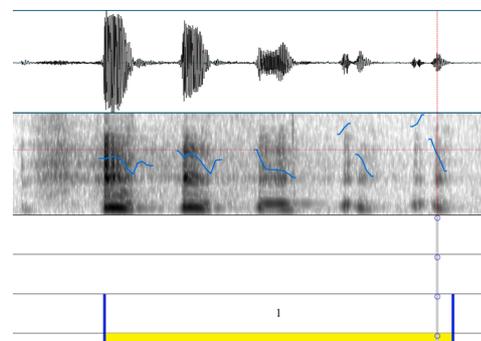


Figure 1 Stereotypical laughter from D-ANS, with alternation of aspiration and vocalic sounds.

Laughter itself is not clearly defined as a discrete activity in common parlance as evidenced by the plethora of terms used in English to describe it – from short shrieks, through snorts, chuckles, and peals. In the scientific literature, there is also a wide range of terminology although there is broad consensus that a ‘stereotypical’ laugh consists of breathy aspirations alternating with vocalic segments of decreasing power (Bachorowski and Owren 1995; Black 1984); an example of such laughter from the D-ANS corpus is shown in Figure 1. However, while this stereotypical laugh is generally produced upon asking an informant to laugh, it has been shown to be only one of several manifestations of laughter present in social interaction, and often not the most prevalent (Trouvain 2003). Also, laughter can occur within speech rather than around it, although these speech laughs, and smiling speech are not regarded as laughter by all researchers. In our investigations we study

laughter in spoken interaction using multimodal corpora of non-scripted (spontaneous) multiparty interaction: the task-oriented AMI meetings corpus (McCowan et al. 2005), and the conversational TableTalk (Jokinen 2009), and D-ANS (Hennig, Chellali, and Campbell 2014) corpora.

1.3 Laughter and topic change

Previous work explored relationships between laughter and topic change in the AMI and TableTalk corpora (Bonin, Campbell, and Vogel 2012), showing greater incidence of laughter in the social TableTalk corpus than in the more task-based AMI corpus. Shared laughter was more prevalent in both corpora than solo laughter. Laughter, and particularly shared laughter, was also seen to be likely immediately before topic change, with a more marked effect in social talk (Gilmartin et al. 2013). These results on multiparty interaction were in line with observations from Conversation Analysis on the distribution of laughter around topic change in two-party conversations (Holt 2010), and strengthened the hypothesis that laughter plays a role in topic transition in social dialogue.

1.4 Obtaining Interaction Data

The information bundle in conversation includes audio (vocal) and visual signals and cues. Vocal information comprises verbal and non-verbal phenomena, including speech, pauses, prosodic contours, voice quality, laughter, coughing, and aspiration noise. The visual channel contains gesture displays, posture variations, and other cues ranging from facial expression, hand and body movements, and eye-gaze to variations in skin colouring (blushing). Some phenomena are tightly linked to the linguistic content of utterances – examples include intonation, beat and iconic gestures, others contribute to the participants' estimation of attitudinal and affective information about each other. In recent years there has been increasing interest in insights into affective and cognitive states gleaned from biosignals – data collected from the body including heart rate, electrodermal activity (skin conductance), skin temperature, pupil

dilation and blood pressure. With technological advances, such measurements can be made unobtrusively over prolonged periods of time using wearable sensors and monitors.

The various facets of conversational exchange are often studied through corpus collection and analysis, with recordings capturing as much as possible of the signal bundle through multiple channels, including audio, video, motion-capture, biosignal data, and data collected from technology used during the interaction such as mice, keyboards or shared whiteboards. Earlier linguistic fieldwork relied on careful written transcription and, later, audio recordings of speech, where a phrase or structure was often elicited directly from an informant in speech community of interest. These elicited pieces of language informed surveys of phonological, syntactic, and semantic aspects of a language or dialect. To gather data on spoken interaction, speech has been elicited by having subjects perform a task or take part in interactions with a clear short-term goal; these 'task-based' interactions are often information gap activities completed by participants exchanging verbal information. Recordings of these tasks then form corpora on which study of dialogue is based. Typical tasks include describing a route through a map as in the HCRC MapTask corpus (Anderson et al. 1991), spotting differences in two pictures as in the DiaPix task in the LUCID (Baker and Hazan 2010) and Wildcat (Van Engen et al. 2010) corpora, ranking items on a list – for example to decide which items would be useful in an emergency (Vinciarelli et al. 2012), and participating in real or staged business meetings as in the ICSI and AMI corpora (Janin et al. 2003; McCowan et al. 2005). While these data collection paradigms do result in corpora of non-scripted dialogue of great utility to researchers, the participants' motivation for taking part in an artificial task is not clear, and the activity is removed from a natural everyday context. It is not certain that tasks such as these can be used to make generalizations about natural conversation (Lemke 2012).

Conversational data has been collected by researchers in the ethnomethodological and conversational analysis traditions by recording everyday home or work situations. Early studies used recordings of telephone calls, as in the suicide hotline data collected by Sacks, Schelgoff's emergency services data, or the conversational data collected by Jefferson (c.f. for example Sacks, Schegloff, & Jefferson, 1974). More domain independent natural telephonic speech data has also been gathered by recording large numbers of real phone conversations, as in the Switchboard corpus (Godfrey, Holliman, and McDaniel 1992), and the ESP-C collection of Japanese telephone conversations (Campbell 2007). Audio corpora of non-telephonic spoken interaction include the Santa Barbara Corpus (DuBois et al. 2000), sections of the ICE corpora (Greenbaum 1991) and of the British National Corpus (BNC-Consortium 2000). Researchers have focused on everyday 'chat' by having subjects wear a microphone throughout the course of their daily lives for extended periods (Campbell 2004), while others have amassed collections of recordings of different types of human activity as in the Gothenburg Corpus (Allwood et al. 2000) which also contains video recordings. With increasing access to video recording and motion capture equipment, and awareness of the multimodal nature of human spoken interaction, rich multimodal corpora are appearing comprising naturalistic encounters with no prescribed task or subject of discussion imposed on participants. These include collections of free-talk meetings, or 'first encounters' between strangers as in the Swedish Spontal, and the NOMCO and MOMCO Danish and Maltese corpora (Edlund et al. 2010; Paggio et al. 2010). In our work, we have studied laughter in task-based and social talk using multimodal multiparty recordings - (the AMI meetings corpus) and social talk (the TableTalk corpus). For our current explorations of laughter in conversation we use the D-ANS corpus of multiparty conversation. Earlier work on laughter in multimodal corpora had brought to light a number of issues in existing laughter annotations. These issues were explored by creating an

experimental annotation scheme for D-ANS. Below we outline the D-ANS corpus in general and the session of interest to our current work, and then describe the experimental laughter annotation procedure we used on D-ANS.

2 Data collection and annotation – the D-ANS corpus

The Dublin Autonomous Nervous System (D-ANS) corpus comprises three sessions of informal English conversation recorded over three days in a living-room like setting with sofa and armchair.

Participant	Sex	Age	Origin
F1	F	30s	USA
F2	F	30s	France
M1	M	30s	Ireland
M2	M	50s	Ireland
M3	M	60s	UK

Table 1 Participants in the D-ANS corpus

There were five participants in total – three men and two women, as shown in Table 1. Two of the men were Irish native English speakers, while the third was a British native English speaker. One woman was a North American native English speaker while the other woman was a French national who had lived in Ireland for some years. She worked in English and her spoken and written command of the language was at the C1 or near-native level on the Common European Scale (Little 2006). Participants were free to speak about any topic they liked and to move around as they wished.

Session 1 consisted of informal chat between the American woman (F1) and the British man (M3). Session 2 contains three hours of conversation between the American woman (F1), the British



Figure 2 Camera Angles in session 3 of D-ANS

man (M3), and one of the Irish men (M2). Session 3, the focus of the analysis described below, contains informal chat between the American woman (F1), one of the Irish men (M1), and the French woman (F2). The camera setup for this session is shown in Figure 2. There was a camera focused on the sofa giving a clear view of speaker F1 sitting on left of sofa and speaker M1 sitting on the right of the sofa. Another camera gave a clear view of speaker F2 sitting on the armchair. Sound was recorded using microphones close to each speaker with an additional omnidirectional microphone on the coffee table between them. The corpus includes measurements of participants' electro-dermal activity (EDA), also known as skin conductance or galvanic skin response. This is a measure of perspiration in the skin, which is linked to arousal. All participants wore biological sensors (Q-sensors) on the underside of each wrist to measure EDA. Q-sensors (Poh, Swenson, and Picard 2010) are wristbands which unobtrusively measure electro-dermal activity (galvanic skin response) in the wrist as well as three degrees of

acceleration. While the wrist, in comparison to the palm or fingers, is a less sensitive location for recording EDA, these sensors were selected to allow free hand movement for gesturing during conversation.

2.2 Annotation of laughter in D-ANS

Several corpora we use have been annotated previously for laughter. The use of existing annotations is attractive to researchers, but in the early stages of our laughter work it became apparent that existing laughter annotation in corpora of interest was not adequate for detailed study, as we encountered several problems including mixtures of point and interval annotation, laughter annotated on the transcription tier at insufficient granularity – e.g. segmented only to the utterance level rather than to word level, and no method for dealing with laughter when it co-occurs with speech. Our observations are in line with problems with laughter annotation outlined by other researchers (Truong and Trouvain 2012). Many of these problems stemmed from the fact that the original

annotation was not created with laughter studies in mind, and so laughter was often annotated as an ‘extra’ or non-verbal phenomenon, often by including a symbol for laughter in the transcription – thus indicating the presence of laughter but not giving information on when exactly it occurred. There is also the lack of agreement in the literature as to what should be considered as laughter. Therefore in our preliminary studies our annotation scheme included only one label for laughter, and annotators were instructed to mark any incidence of laughter that they noticed.

We addressed these problems in our earlier work on TableTalk by performing a new manual laughter annotation using ELAN (Wittenburg et al. 2006) with separate laugh tracks for each speaker, annotating laughter according to the MUMIN scheme (Allwood et al. 2004). During annotation we noticed that many laughs were not acoustically sounded, (a requirement in the MUMIN definition), or too quiet to be picked up by microphone, and therefore we performed extra passes over the data using Praat with the sound files and Elan with video files to capture as much of the laughter as possible. This was a time-consuming process and raised the question of how much laughter was solely identifiable by video or audio alone. To investigate this question we devised an experimental annotation scheme by adding sounded (acoustic) and silent (non-acoustic) laugh tracks. We used this scheme with D-ANS to investigate silent and sounded perception of laughter. Annotation of laughter in D-ANS was performed in three passes – video only, audio only, and video with sound.

2.3 Video only ‘silent’ annotation

The video annotation was performed by two annotators, one ‘naïve’ annotator, a member of the public who was not involved in the linguistics or communications field. The second annotator was a speech researcher. For the video only (‘silent’) passes the two annotators were provided with silent video of the data and asked to mark intervals where they saw laughter. The annotation was performed in ELAN. For each

video, a linked annotation file was created containing a single tier in which to annotate laughter for one of the speakers in the video. The video was marked for laughter while annotators watched in real time. ‘Hot keys’ were set up in the annotation file so that annotators could press a key when they saw laughter and press again when the laughter ended. Participants were allowed to pause the video, but were discouraged from replaying sections and changing their markings except in cases where they recognized that they had forgotten to mark laughter endpoints. In real time annotation of this type, there is an issue of lag or reaction time in the button pressing by participants. While Elan does offer a facility to automatically factor in a lag time correction to annotations, this was not used as it would correct for a constant lag. In real time annotation, it seems more likely that the initial delays in reaction would be greater than those later in the process as the annotator became accustomed to the task and to the speaker they were annotating. To address this contingency, before annotating a particular speaker, annotators were given ‘practice’ ELAN files containing video of the speaker taken from a different section of the corpus. It was hoped that this would allow any lag to settle before the annotator started on the video of interest. In any case, the real time laughter annotation by naïve annotators was not regarded as a highly temporally exact segmentation of laughter but rather as an indication of the occurrence of laughter, and was used as such in the analysis.

2.4 Audio only ‘sounded’ annotation and standard annotation

The audio only annotation was performed in Praat (Boersma and Weenink 2010) by two annotators marking sound recordings of the data with intervals where they heard laughter. The third annotation was a standard annotation made using the video and audio tracks together in ELAN. We performed two analyses on the D-ANS laughter data – investigating laughter around topic change and exploring differences in silent video only and sounded audio only annotations of laughter.

P	C	A1	A2	Agreed	Agreed %
M1	A	51	52	48	87
F1	A	73	72	70	93
F2	A	52	50	49	92
M1	V	98	90	86	86
F1	V	69	70	64	85
F2	V	105	94	92	86

Table 2 Agreement between Annotators (A1 and A2) on Audio only (A) and Video only (V) annotations.

3 Results

In the data we noted that many laugh annotations were separated by a short silence, often while the participant was breathing, so for the purpose of our analyses we define a ‘laugh event’ in which we amalgamated annotations of laughter which were separated by less than 1 second.

3.1 Multimodal annotation of laughter

We analysed the relationship between the silent and sounded laughter annotations in categorical terms by looking at raters’ agreement on the incidence rather than the duration of laughter. We counted the number of laugh events noted in each of the Audio only (sounded) and Video only (silent) annotations (Table 2). We found inter-rater agreement to be high for each condition between annotators annotating the same modality. In the silent video condition agreement between the annotators ranged from 85 to 86% depending on speaker, while for the audio only annotations agreement ranged from 87 to 93% agreement per speaker. We then discarded all cases of audio or video laugh events which were marked by only one annotator, leaving a dataset with only the ‘agreed’ laughs - laughter recorded by both annotators in a particular modality (Table 3).

P	A and V	A only	V only
M1	37	3	44
F1	56	7	5
F2	45	2	47

Table 3 ‘Agreed’ laugh events, where all annotators in the modality recorded laughs

Table 3 shows the resulting per-speaker counts for agreed laughs appearing in the annotations for the three speakers F1, F2, and M2 as described above. The ‘A and V’ column shows laughs which were recorded by both the Audio only annotators and the Video only (silent video) annotators. The ‘A only’ column shows laughs only picked up by Audio only annotators, while the ‘V only’ column shows laughs picked up by silent video annotators but not by audio annotators.

We found that most cases where annotations were made on video but not on the audio (V only) involve a combination of head tilting (pitch) and a growing broad smile or wide or toothy grin. In annotations on the audio but not the video (A only), most involve laughter co-occurring with speech with a much smaller number of cases where the annotation was of a short phrase initial or final laugh or snort.

3.2 Laughter and topic change

We extended previous analyses of shared and solo laughter in relation to topic change to the D-ANS data in order to investigate whether earlier results on the likelihood of laughter in the vicinity of topic change (Bonin, Campbell, and Vogel 2012; Gilmartin et al. 2013) would generalize to the D-ANS corpus.

In Session 3 of D-ANS, there were a total of 80 shared laughter events and 21 topics discussed. The distance from each of the 20 topic change points to the last shared laugh, as shown in Figure 3, ranged from 0 to 10.2 seconds, with 90%



Figure 3 Histogram of distances from topic changes to the endpoint of the last shared laugh

of topic changes occurring within 5 seconds of the end of an interval of shared laughter. Modelling the situation as a binomial distribution with the probability of any point falling within 5 seconds of the end of a shared laugh equal to the ratio of SL to T, where SL is the number of seconds on the recording meeting the criterion of being within 5 seconds of the end point of a shared laugh event and T is the total number of seconds in the recording, we can reject the null hypothesis that topic change points are randomly placed with respect to shared laughter ($p < 0.001$).

4 Discussion

The results of the topic transition analysis on a section of the DANS corpus show more shared than solo laughter in multiparty social dialogue; in line with earlier results and reports in the literature on the social nature of laughter. The strong likelihood of laughter before topic change points found in our analysis of D-ANS echoes the results of earlier work on TableTalk and AMI. A possible explanation for this tendency for topic changes to occur in or soon after shared laughter stems from the fact that in social talk there is set structure to follow and no agenda of topics lined up as there is in a meeting, nor are roles predetermined (Cheepen 1988; Laver 1975; Thornbury and Slade 2006). Without set roles or tasks or indeed a chairperson to ‘move things along’, the management of topic falls back on the participants. In extended social talk, conversation evolves as a series of often narrative longer ‘chunks’ interspersed with short ‘chat’

exchanges. Thus participants share the burden of providing entertaining speech input to the group. As topics are exhausted, there is a danger of uncomfortable silence, and participants may avoid this by entering a buffer of shared laughter, reinforcing positive bonds, and providing space for a new topic to be produced and the baton of speaker passed to another participant. Laughter may thus function around topic change much as the ‘idling’ behavior noted in social talk, when there is nothing much to be said but the impetus seems to be to keep the conversational ball in the air (Schneider 1988).

The pilot study results on multimodality indicate that careful annotation on the audio channel picks up most stereotypical sounded laughter, but can result in false positives in the case of speech laughs, although this phenomenon was observed in only one of the three speakers examined. Naïve human annotators asked to mark laughter watching silent video picked up the vast bulk of audio laughter, but also identified smiles and head nods accompanied by a wide grin as laughter. While this could be viewed as misclassification, it happens on a large enough scale to beg the question of whether such behavior should be regarded as silent laughter and thus may point to the need for a clearer taxonomy of laughter. In terms of applications of our findings on multimodal aspects of laughter, video based automatic identification of laughter may be an attractive prospect. Automatic identification of laughter on the audio stream is possible for stereotypical laughter (Scherer et al. 2009) but requires clean audio signals from close-coupled microphones. This is a limitation for real-world use of audio-based technology for laughter detection. Video signals are more robust, and identification on video data is an attractive idea, however there is a need for clear definitions of the various phenomena identified as laughter outside of the narrow stereotypical description. During our work on laughter in conversational corpora we have noted the need to re-annotate, and then expand our annotation scheme in view of observations during manual annotation. While data annotation is time-consuming and labour-

intensive work, it is invaluable for a fuller understanding of the dynamics of human interaction. Indeed, close examination of data has revealed subtleties that may have been missed had we simply used pre-existing annotations.

5 Future work

We are currently investigating the interplay of laughter and biosignals in D-ANS. We are particularly interested in any variation in ‘task’ and ‘chat’ dialogue, in terms of laughter and of measured electro-dermal activity (EDA). EDA has been linked to levels of emotional arousal (Dawson, Schell, and Fillion 2007) and to cognitive load (Shi et al. 2007), with work in psychology observing lower cognitive load in social chat than in talk arising during tasks (Kahneman 2011). Laughter has been observed to be more frequent in social than in task-based dialogue, and to be active around topic changes. This knowledge may help distinguish whether stretches of conversation are transactional or social in nature. It may thus be possible to contribute towards technological extraction of important or content-rich sections of dialogue using insights gained in our work. To further investigate the multimodality of laughter, we are creating more detailed laughter annotations which will allow us to further explore whether the silent phenomena our naïve annotators marked as laughter are functionally different to sounded laughter in terms of where and in what capacity they occur in conversation – in the listener/speaker, as backchannels, before or after audio laughter, in solo or shared laughter. We are currently extending our investigations to the D64 corpus of conversational English to test the generalizability of our findings.

Acknowledgments

This work is supported by the Fastnet Project – Focus on Action in Social Talk: Network Enabling Technology funded by Science Foundation Ireland (SFI) 09/IN.1/I2631, and by the Università degli Studi di Genova and the PAVIS department at the Istituto Italiano di Tecnologia.

References

- Allen, James F., Donna K. Byron, Myroslava Dzikovska, George Ferguson, Lucian Galescu, and Amanda Stent. 2001. ‘Toward Conversational Human-Computer Interaction’. *AI Magazine* 22 (4): 27.
- Allwood, J., L. Cerrato, L. Dybkær, and P. Paggio. 2004. ‘The MUMIN Multimodal Coding Scheme’. In *Proc. Workshop on Multimodal Corpora and Annotation*.
- Allwood, Jens, Maria Björnberg, Leif Grönqvist, Elisabeth Ahlsén, and Cajsa Ottesjö. 2000. ‘The Spoken Language Corpus at the Department of Linguistics, Göteborg University’. In *FQS–Forum Qualitative Social Research*. Vol. 1.
- Anderson, A.H., M. Bader, E.G. Bard, E. Boyle, G. Doherty, S. Garrod, S. Isard, et al. 1991. ‘The HCRC Map Task Corpus’. *Language and Speech* 34 (4): 351–66.
- Bachorowski, J. A., and M. J. Owren. 1995. ‘Vocal Expression of Emotion: Acoustic Properties of Speech Are Associated with Emotional Intensity and Context’. *Psychological Science*, 219–24.
- Baker, Rachel, and Valerie Hazan. 2010. ‘LUCID: A Corpus of Spontaneous and Read Clear Speech in British English’. In *Proceedings of the DiSS-LPSS Joint Workshop 2010*.
- Bakhtine, Mikhail Mikhaïlovitch. 1986. *Speech Genres and Other Late Essays*. 8. University of Texas Press.
- Black, Donald W. 1984. ‘Laughter’. *JAMA: The Journal of the American Medical Association* 252 (21): 2995–98.
- BNC-Consortium. 2000. ‘British National Corpus’. URL [Http://www.Hcu.Ox.Ac.uk/BNC](http://www.Hcu.Ox.Ac.uk/BNC).
- Boersma, Paul, and David Weenink. 2010. *Praat: Doing Phonetics by Computer [Computer Program], Version 5.1. 44*.
- Bonin, Francesca, Nick Campbell, and Carl Vogel. 2012. ‘Laughter and Topic Changes: Temporal Distribution and Information Flow’. In *Cognitive Infocommunications (CogInfoCom), 2012 IEEE 3rd International Conference on*, 53–58.
- Burling, R. 2007. *The Talking Ape: How Language Evolved*. Vol. 5. Oxford University Press, USA.

- Campbell, Nick. 2004. 'Speech & Expression; the Value of a Longitudinal Corpus.' In *Proc. LREC 2004*.
- . 2007. 'Approaches to Conversational Speech Rhythm: Speech Activity in Two-Person Telephone Dialogues'. In *Proc XVIth International Congress of the Phonetic Sciences, Saarbrücken, Germany*, 343–48.
- Cheepen, Christine. 1988. *The Predictability of Informal Conversation*. Pinter London.
- Dawson, Michael E., Anne M. Schell, and Diane L. Filion. 2007. 'The Electrodermal System'. *Handbook of Psychophysiology*, 159.
- DuBois, John W., W. L. Chafe, C. Meyer, and S. A. Thompson. 2000. *Santa Barbara Corpus of Spoken American English. CD-ROM. Philadelphia: Linguistic Data Consortium*.
- Dunbar, R. 1998. *Grooming, Gossip, and the Evolution of Language*. Harvard Univ Press.
- Edlund, Jens, Jonas Beskow, Kjell Elenius, Kahl Hellmer, Sofia Strömbergsson, and David House. 2010. 'Spontal: A Swedish Spontaneous Dialogue Corpus of Audio, Video and Motion Capture.' In *LREC*.
- Gilmartin, Emer, Francesca Bonin, Carl Vogel, and Nick Campbell. 2013. 'Laughter and Topic Transition in Multiparty Conversation'. In *Proceedings of the SIGDIAL 2013 Conference*, 304–8. Metz, France: Association for Computational Linguistics.
- Gilmartin, Emer, and Nick Campbell. 2014. 'More Than Just Words: Building a Chatty Robot'. In *Natural Interaction with Robots, Knowbots and Smartphones*.
- Glenn, Phillip J. 2003. *Laughter in Interaction*. Cambridge University Press Cambridge.
- Godfrey, John J., Edward C. Holliman, and Jane McDaniel. 1992. 'SWITCHBOARD: Telephone Speech Corpus for Research and Development'. In *Acoustics, Speech, and Signal Processing, 1992. ICASSP-92., 1992 IEEE International Conference on*, 1:517–20.
- Greenbaum, Sidney. 1991. 'ICE: The International Corpus of English'. *English Today* 28 (7.4): 3–7.
- Hennig, Shannon, Ryad Chellali, and Nick Campbell. 2014. 'The D-ANS Corpus: The Dublin-Autonomous Nervous System Corpus of Biosignal and Multimodal Recordings of Conversational Speech.' In *Proc. LREC 2014, Reykjavik, Iceland*.
- Holt, Elizabeth. 2010. 'The Last Laugh: Shared Laughter and Topic Termination'. *Journal of Pragmatics* 42 (6): 1513–25.
- Janin, Adam, Don Baron, Jane Edwards, Dan Ellis, David Gelbart, Nelson Morgan, Barbara Peskin, Thilo Pfau, Elizabeth Shriberg, and Andreas Stolcke. 2003. 'The ICSI Meeting Corpus'. In *Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP'03). 2003 IEEE International Conference on*, 1:I–364.
- Jokinen, Kristiina. 2009. 'Gaze and Gesture Activity in Communication'. In *Universal Access in Human-Computer Interaction. Intelligent and Ubiquitous Interaction Environments*, 537–46. Springer.
- Kahneman, D. 2011. *Thinking, Fast and Slow*. Farrar Straus & Giroux.
- Laver, John. 1975. 'Communicative Functions of Phatic Communion'. *Organization of Behavior in Face-to-Face Interaction*, 215–38.
- Lemke, Jay L. 2012. 'Analyzing Verbal Data: Principles, Methods, and Problems'. In *Second International Handbook of Science Education*, 1471–84. Springer.
- Little, David. 2006. 'The Common European Framework of Reference for Languages: Content, Purpose, Origin, Reception and Impact'. *Language Teaching* 39 (03): 167–90.
- Malinowski, B. 1923. 'The Problem of Meaning in Primitive Languages'. *Supplementary in the Meaning of Meaning*, 1–84.
- McCowan, Iain, Jean Carletta, W. Kraaij, S. Ashby, S. Bourban, M. Flynn, M. Guillemot, T. Hain, J. Kadlec, and V. Karaiskos. 2005. 'The AMI Meeting Corpus'. In *Proceedings of the 5th International Conference on Methods and Techniques in Behavioral Research*. Vol. 88.
- Mehu, Marc, and Robin IM Dunbar. 2008. 'Relationship between Smiling and Laughter in Humans (Homo Sapiens): Testing the Power Asymmetry Hypothesis'. *Folia Primatologica* 79 (5): 269–80.
- Paggio, Patrizia, Jens Allwood, Elisabeth Ahlsén, and Kristiina Jokinen. 2010. 'The

- NOMCO Multimodal Nordic Resource—goals and Characteristics’.
- Poh, M. Z., N. C. Swenson, and R. W. Picard. 2010. ‘A Wearable Sensor for Unobtrusive, Long-Term Assessment of Electrodermal Activity’. *Biomedical Engineering, IEEE Transactions on* 57 (5): 1243–52.
- Provine, Robert R. 1993. ‘Laughter Punctuates Speech: Linguistic, Social and Gender Contexts of Laughter’. *Ethology* 95 (4): 291–98.
- . 2004. ‘Laughing, Tickling, and the Evolution of Speech and Self’. *Current Directions in Psychological Science* 13 (6): 215–18.
- Sacks, H., E.A. Schegloff, and G. Jefferson. 1974. ‘A Simplest Systematics for the Organization of Turn-Taking for Conversation’. *Language*, 696–735.
- Scherer, Stefan, Friedhelm Schwenker, Nick Campbell, and Günther Palm. 2009. ‘Multimodal Laughter Detection in Natural Discourses’. In *Human Centered Robot Systems*, 111–20. Springer.
- Schneider, Klaus P. 1988. *Small Talk: Analysing Phatic Discourse*. Vol. 1. Hitzeroth Marburg.
- Shi, Yu, Natalie Ruiz, Ronnie Taib, Eric Choi, and Fang Chen. 2007. ‘Galvanic Skin Response (GSR) as an Index of Cognitive Load’. In *CHI’07 Extended Abstracts on Human Factors in Computing Systems*, 2651–56.
- Thornbury, Scott, and Diana Slade. 2006. *Conversation: From Description to Pedagogy*. Cambridge University Press.
- Trouvain, Jürgen. 2003. ‘Segmenting Phonetic Units in Laughter’. In *Proceedings of the 15th International Congress of Phonetic Sciences. Barcelona: Universitat Autònoma de Barcelona*, 2793–96.
- Truong, Khiet P., and Jurgen Trouvain. 2012. ‘Laughter Annotations in Conversational Speech Corpora—Possibilities and Limitations for Phonetic Analysis’. *Proceedings of the 4th International Workshop on Corpora for Research on Emotion Sentiment and Social Signals*, 20–24.
- Van Engen, Kristin J., Melissa Baese-Berk, Rachel E. Baker, Arim Choi, Midam Kim, and Ann R. Bradlow. 2010. ‘The Wildcat Corpus of Native-and Foreign-Accented English: Communicative Efficiency across Conversational Dyads with Varying Language Alignment Profiles’. *Language and Speech* 53 (4): 510–40.
- Vinciarelli, Alessandro, Hugues Salamin, Anna Polychroniou, Gelareh Mohammadi, and Antonio Origlia. 2012. ‘From Nonverbal Cues to Perception: Personality and Social Attractiveness’. In *Cognitive Behavioural Systems*, 60–72. Springer.
- Wittenburg, Peter, Hennie Brugman, Albert Russel, Alex Klassmann, and Han Sloetjes. 2006. ‘Elan: A Professional Framework for Multimodality Research’. In *Proceedings of LREC*. Vol. 2006.